



Wilson, J., Barker, S., & Ridgwell, A. J. (2012). Assessment of the spatial variability in particulate organic matter and mineral sinking fluxes in the ocean interior:: Implications for the ballast hypothesis. *Global Biogeochemical Cycles*, 26(04), [GB4011]. 10.1029/2012GB004398

Publisher's PDF, also known as Final Published Version

Link to published version (if available):
[10.1029/2012GB004398](https://doi.org/10.1029/2012GB004398)

[Link to publication record in Explore Bristol Research](#)
PDF-document

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:
<http://www.bristol.ac.uk/pure/about/ebr-terms.html>

Take down policy

Explore Bristol Research is a digital archive and the intention is that deposited content should not be removed. However, if you believe that this version of the work breaches copyright law please contact open-access@bristol.ac.uk and include the following information in your message:

- Your contact details
- Bibliographic details for the item, including a URL
- An outline of the nature of the complaint

On receipt of your message the Open Access Team will immediately investigate your claim, make an initial judgement of the validity of the claim and, where appropriate, withdraw the item in question from public view.

Assessment of the spatial variability in particulate organic matter and mineral sinking fluxes in the ocean interior: Implications for the ballast hypothesis

J. D. Wilson,¹ S. Barker,¹ and A. Ridgwell²

Received 30 April 2012; revised 10 August 2012; accepted 16 September 2012; published 2 November 2012.

[1] Multiple linear regression analysis (MLRA) applied to sediment trap data has been highly influential in identifying a plausible ‘ballasting’ mechanism that directly links the settling fluxes of particulate organic carbon (POC) to those of denser, inorganic minerals. However, analysis to date has primarily been carried out at the global scale, missing spatial variability in the flux relationships that may be important. In this paper, Geographically Weighted Regression (GWR) is applied to an updated deep (>1500 m) sediment trap database ($n = 156$), using the MLRA approach of Klaas and Archer (2002) but now allowing the carrying coefficients to vary in space. While the global mean carrying coefficient values for CaCO_3 , opal, and lithogenics are broadly consistent with previous work, the GWR analysis reveals the existence of substantial and statistically significant spatial variability in all three carrying coefficients. In particular, the absence of a strong globally uniform relationship between CaCO_3 and POC in our spatial analysis calls into question whether a simple ballasting mechanism exists. Instead, the existence of coherent spatial patterns in carrying coefficients, which are reminiscent of biogeochemical provinces, points toward differences in specific pelagic ecosystem characteristics as being the likely underlying cause of the flux relationships sampled by sediment traps. Our findings present a challenge to ocean carbon cycle modelers who to date have applied a single statistical global relationship in their carbon flux parameterizations when considering mineral ballasting, and provide a further clue as to how the efficiency of the biological pump in the modern ocean is regulated.

Citation: Wilson, J. D., S. Barker, and A. Ridgwell (2012), Assessment of the spatial variability in particulate organic matter and mineral sinking fluxes in the ocean interior: Implications for the ballast hypothesis, *Global Biogeochem. Cycles*, 26, GB4011, doi:10.1029/2012GB004398.

1. Introduction

[2] Sinking particles transfer particular organic carbon (POC) and associated nutrients from the upper ocean to the deep ocean and sediments in a process known as the biological pump [Honjo *et al.*, 2008]. These particles, ultimately derived from the growth of phytoplankton at the sunlit surface and carbon fixation through photosynthesis, are also often associated with biominerals such as biogenic silica (opal) and calcium carbonate (CaCO_3). As these particles sink, the majority of POC and associated nutrients are remineralized (predominantly) by bacterial metabolic processes and zooplankton flux feeding in the upper ~ 1000 m, leaving a small (5–10%) fraction (relative to that at 100 m) sinking to depth [Stemmann *et al.*, 2004; Loubere *et al.*,

2007; Honjo *et al.*, 2008]. Understanding the processes that control the efficiency of the biological pump in transporting carbon and nutrients to depth is key to understanding how the marine carbon cycle functions and regulates atmospheric carbon dioxide (CO_2) (e.g. Archer and Maier-Reimer [1994]).

[3] The ratio of particulate inorganic to organic carbon (PIC:POC) within sinking particles is known as the ‘rain ratio’ and is important in communicating changes at the surface to the deep ocean and sediments. For instance, on time-scales of a few thousand years, a reduction in the export rain ratio of 40%, if communicated to the sediments could, in theory, lead to a 70–90 ppm drawdown of atmospheric CO_2 through the increased dissolution of carbonate sediments [Archer and Maier-Reimer, 1994]. However, studies based on the analysis of deep sediment trap data have observed a strong global correlation between mass fluxes of POC and CaCO_3 , suggesting some mechanism of coupling exists between these important parameters at depth [Armstrong *et al.*, 2001]. The ‘ballast hypothesis’ posits that CaCO_3 , and to a lesser extent opal and lithogenic material, aids the sinking of particles through increasing mean aggregate density

¹School of Earth and Ocean Sciences, Cardiff University, Cardiff, UK.

²School of Geographical Sciences, University of Bristol, Bristol, UK.

Corresponding author: J. D. Wilson, School of Earth and Ocean Sciences, Cardiff University, Main Building, Park Place, Cardiff CF10 3AT, UK. (wilsonjd@cardiff.ac.uk)

[Armstrong *et al.*, 2001; Klaas and Archer, 2002]. If true, this would have the effect of buffering changes in the rain ratio originating at the surface and reducing the potential for altering atmospheric CO₂ [Ridgwell, 2003]. In the context of ocean acidification and the potential for decreased pelagic calcification driven by falling surface ocean carbonate saturation, ballasting creates a positive feedback to CO₂ by reducing the efficiency of the biological pump [Barker *et al.*, 2003; Heinze, 2004; Riebesell *et al.*, 2009].

[4] Particles have been shown to sink faster under laboratory conditions due to the relatively high density of CaCO₃ [Ploug *et al.*, 2008; Engel *et al.*, 2009; Iversen and Ploug, 2010], supporting the statistically-based ballast hypothesis. However, other lines of evidence point to alternative interpretations. For instance, rolling tank experiments showed the POC:mineral ratio was dependent on the amount of POC present acting as a ‘glue’, aggregating mineral fluxes [Passow, 2004; Passow and De La Rocha, 2006; De La Rocha *et al.*, 2008] and controlling the sinking of CaCO₃ rather than vice versa. Alternatively, variations in surface ecosystem composition might dictate the packaging and remineralization of particles [Francois *et al.*, 2002; Lam and Bishop, 2007; Lam *et al.*, 2011] and give rise to the flux relationships observed at greater depth. These alternative explanations significantly challenge our mechanistic understanding of the dynamics of POC fluxes and create substantial uncertainty in both the magnitude and sign of carbon cycle feedbacks to possible future perturbations [Barker *et al.*, 2003; Riebesell *et al.*, 2009].

[5] The global sediment trap analysis of Klaas and Archer [2002] has been highly influential in quantifying the correlation between POC and CaCO₃ and helping to formulate the ballasting hypothesis. In that study, the mass flux of POC was expressed as a linear function of three dominant mineral fluxes (CaCO₃, opal and lithogenic material), using multiple linear regression analysis (MLRA). The derived ‘carrying coefficients’ (the regression coefficients) were largest for CaCO₃ (0.070–0.094), lowest for opal (0.023–0.030), and rather variable for lithogenics (0.035–0.071) [Klaas and Archer, 2002] with the resulting statistical models able to explain a large proportion of the observed variability in POC flux. The three mineral model provides a basis for understanding global variability in POC to mineral ratios and can replace this term in the mechanistic model of Armstrong *et al.* [2001], making this a useful method for parameterizing particle fluxes in a range of ocean carbon cycle models [Howard *et al.*, 2006; Oka *et al.*, 2008; Hofmann and Schellnhuber, 2009].

[6] The underlying assumption when analyzing the global database in this way is that the statistical relationships (the coefficient values) are the same for any location in the ocean, i.e. they are assumed stationary in space. The use of these global statistical relationships in models then explicitly makes this same assumption. However, it is reasonable to expect that these relationships may not be constant in space (or time) i.e. they may exhibit spatial nonstationarity, which can be characterized by a non-random distribution of residuals in space [Fotheringham *et al.*, 1998]. This potentially raises issues for the interpretation of global regression coefficients and their explicit use as a parameterization in modeling studies.

[7] Spatial variability in the relationship between POC and minerals was noted in sediment trap data by Ragueneau *et al.* [2006] and De La Rocha and Passow [2007] who suggested global MLRA was, therefore, inappropriate and may have misleadingly resulted in the low carrying coefficients obtained for opal and lithogenics. Boyd and Trull [2007] also note that using global annual mean fluxes ignores a large part of variability resulting from processes like El Niño as well as the biogeochemical sources of the fluxes. A previous basin-scale analysis (Table 1) showed considerable regional variability in the dominance of one mineral over another. Global MLRA may then be hiding important regional variability which has implications for how the ballast mechanism is interpreted and particularly for how it is mechanistically implemented in global models.

[8] To date there has been no general assessment of the spatial variability of the carrying coefficients of ballast minerals. This is important for understanding the previously observed differences between global coefficients and those seen from individual sites (Table 1). Ragueneau *et al.* [2006] took the first step in this respect and applied MLRA to sediment trap data divided by major ocean basin. This broad delineation was reasonably justified but further reduction of the spatial scale poses particular problems. Smaller spatial groupings for regression could be justified, such as biogeochemical provinces (see Vichi *et al.* [2011]) but this introduces a level of subjectivity, as well as problems with the relatively sparse sampling coverage of sediment trap data sets compared to the number of biogeographical provinces. In response to this we describe the novel application of Geographically Weighted Regression (GWR), that allows coefficients to vary in space and helps avoid the problems stated above. The data set and technique are described in section 2 and applied using the carrying coefficient approach of Klaas and Archer [2002] to explore the spatial variability of these statistical parameters.

2. Methodology

2.1. Sediment Trap Data

[9] An updated global sediment trap data set has been collated for this study. The flux and metadata from sediment traps are available in the auxiliary material.¹ The majority of the data set used here is from the U.S. Joint Global Ocean Flux Study (JGOFS) available online via: <http://usjgofs.whoi.edu/mzweb/syn-mod.htm>. A full description of the JGOFS data set and methodologies can be found in Honjo *et al.* [2008]. Other data sets were obtained from additional studies and the World Data Centre for Environmental Sciences (WDC-MARE) online database.

[10] Sediment traps at relatively shallow depths (approx. <1000–1500 m) have been shown to be inefficient at trapping particulate material [Scholten *et al.*, 2001; Yu *et al.*, 2001]. For this reason, and to be consistent with the bulk of previous work, only flux data at >1500 m were primarily considered here, although additional analysis with a >1000 m cut-off depth was carried out to enable comparison to Ragueneau *et al.* [2006]. Selected data includes only sediment traps that sampled over a minimum period of 320 days

¹Auxiliary materials are available in the HTML. doi:10.1029/2012GB004398.

Table 1. Carrying Coefficients (i.e., *Klaas and Archer* [2002]) Derived From Previous Multiple Linear Regression Analyses Applied to a Range of Global, Regional and Time Series Sediment Trap Datasets^a

Global Annual Average		CaCO ₃	Opal	Lithogenic ^b	R ²
<i>Klaas and Archer</i> [2002] ^c	n = 78	0.075	0.029	0.052	–
	>2000 m	0.064–0.086	0.020–0.037	0.034–0.070	
<i>Francois et al.</i> [2002] ^c	n = 62	0.074	0.015	0.074	0.93
	>2000 m	0.064–0.084	0.08–0.022	0.051–0.097	
<i>Ragueneau et al.</i> [2006]	n = 189	0.081	0.031	0.035	0.89
	>1000 m	0.073–0.089	0.023–0.039	0.029–0.041	
Regional Annual Average					
<i>Ragueneau et al.</i> [2006] (Atlantic)	n = 84 >1000 m	0.077 0.053–0.101	0.171 0.047–0.186	0.031 0.023–0.039	0.87
<i>Ragueneau et al.</i> [2006] (Indian)	n = 16 >1000 m	0.026 ^d –0.007–0.057	0.201 0.123–0.279	0.015 ^d –0.079–0.049	0.96
<i>Ragueneau et al.</i> [2006] (Pacific)	n = 89 >1000 m	0.063 0.055–0.071	0.041 0.035–0.046	0.024 0.018–0.030	0.95
Regional Time series					
<i>Wong et al.</i> [1999] ^c (Ocean Station P)	1982–1993 3800 m	0.021 0.002–0.039	0.013 0.008–0.034	0.0233 0.170–0.297	0.69
<i>Conte et al.</i> [2001] ^c (Bermuda SCIFF)	1978–1984 3200 m	0.045 0.038–0.053	0.063 0.024–0.102	0.065 0.034–0.096	0.98
<i>Honda and Watanabe</i> [2010] ^c (W. Pacific Subarctic Gyre)	1998–2006 4810	0.025 n/a	0.044 n/a	–0.006 n/a	0.92

^aRanges given indicate 95% range of carrying coefficients (2× standard error).^bAll lithogenic material is estimated as $Total\ Mass - (POC * POM\ Conversion\ factor) + CaCO_3 + Opal$, except *Honda and Watanabe* [2010] which was derived from AI measurements.^cAdapted from Table 3a in *Boyd and Trull* [2007].^dValues from multiple regression analysis were insignificant at $p > 0.05$.^eData were normalized to average of each time series component before regression analysis.

to maximize the quantity of data and its spatial coverage while retaining a reasonable annual coverage. Finally, data were excluded if any observations were missing for major components (total mass flux, POC, PIC, biogenic silica). The mass fluxes of CaCO₃ and opal were estimated from PIC and biogenic Si using conversion factors of 8.33 and 2.14 respectively [*Mortlock and Froelich*, 1989]. The flux of lithogenic material was estimated as the remaining fraction of total mass flux when CaCO₃, opal and particulate organic matter are subtracted. In a small number of cases this produced negative values for lithogenic flux, which were then treated as zero (as in *Salter et al.* [2010]).

[11] The resulting data set comprises 156 individual sediment trap observations which include data on POC, PIC (as CaCO₃), biogenic silica (as opal), lithogenic and total mass fluxes. The data set includes observations from 25 biogeochemical provinces. In comparison to the previous global data sets [*Francois et al.*, 2002; *Klaas and Archer*, 2002], this data set is larger in size ($n = 156$ c.f. $n = 62$ –78) and provides greater spatial coverage, particularly for the southern hemisphere. The data set is of a comparable size to a recent sediment trap collation [*Honjo et al.*, 2008]. However, because approximately half of the data set is in common with previous analyses, we would not expect the results of our global analysis to be substantially different from previous studies.

2.2. Regression Analysis

2.2.1. Global Regression Model

[12] Here we apply the multiple linear regression analysis used in *Klaas and Archer* [2002]. The basic regression analysis expresses the flux of POC at depth (F_{POC}) as a

function of the fluxes of CaCO₃ (F_{CaCO_3}), Opal (F_{Opal}) and lithogenic material (F_{litho}) at depth (z):

$$F_{POC}(z) = \beta_0 + \beta_{CaCO_3} \cdot F_{CaCO_3}(z) + \beta_{Opal} \cdot F_{Opal}(z) + \beta_{litho} \cdot F_{litho}(z) \quad (1)$$

Previous analyses assumed that the regression passed through the origin [*Ragueneau et al.*, 2006] requiring that the flux of POC must be zero when the flux of minerals is zero. We include an additional intercept term (β_0) so that the analysis is amenable and directly comparable to the geographically weighted technique we employ (Section 2.2.2) and also to create a more general model in which it is possible that there could additional POC not directly associated with the mineral flux.

2.2.2. Geographically Weighted Regression Model

[13] Geographically weighted regression is a relatively novel but simple technique of regression which allows the estimation of local statistical parameters (for a full description see *Fotheringham et al.* [2002], also *Brunsdon et al.* [1998] and *Fotheringham et al.* [1998]). The global multiple linear regression model can be considered as:

$$y_i = \alpha_0 + \sum_k a_k x_{ik} + \epsilon_i, \quad (2)$$

where k predictors are used to predict y at the i_{th} point in space. The global model can be re-written to estimate local parameters as:

$$y_i = \alpha_0(u_i, v_i) + \sum_k a_k(u_i, v_i) x_{ik} + \epsilon_i, \quad (3)$$

where (u_i, v_i) denotes the coordinates of the i th point in space. $a_k(u_i, v_i)$ is then a realization of the continuous function $a_k(u, v)$ at point i [Fotheringham *et al.*, 1998]. The global model can be seen as a case where the parameter surface is considered constant in space. GWR approximates the above equation by selecting a subset of data around i which is weighted according to their distance from i . This assumes that parameters display a degree of spatial consistency such that parameters become increasingly different as distance increases from i .

[14] The weighting function (kernel) used in GWR can take different forms. The simplest approach applies a weight of 1 to all data within a set distance (d) of i and 0 to data outside of this area. This function creates artificial boundaries and could create artifacts in the patterns of parameters estimated. An alternative is to use a continuous weighting function such as in exponential form or a bi-square function that approximates a Gaussian weighting within the bandwidth and zero beyond this (4):

$$w_{ij} = \left[1 - (d_{ij}/\beta)^2\right]^2 \text{ if } d_{ij} < \beta = 0 \text{ otherwise,} \quad (4)$$

where the weight applied (w_{ij}) is a function of distance (d) between i and j and a parameter referred to as the bandwidth (β), defining the total distance of the subset of data [Fotheringham *et al.*, 2002]. The bandwidth can be defined either as a fixed distance (fixed kernel) or as a number of nearest neighbors (adaptive kernel) in the data set and is a global parameter. The latter definition allows the weighting function to respond to changes in sampling density where a fixed kernel may be unsuitable to apply. In both definitions the weighting function defines a ‘bump of influence’ around data at point i .

[15] The size of the bandwidth is a critical factor in GWR as it defines the area of influence of each regression. A larger bandwidth will lose information about the spatial variability in coefficients and bias the results toward the global regression. If regression coefficients are considered to be a continuous field in space, then each data point in the subset of data will have a unique coefficient value, but defining a subset of data around each point forces the coefficient to a common (essentially an average) value for that subset. Therefore, coefficients can never be completely unbiased because there is always a level of spatial averaging. To minimize the bias of coefficients, a small subset of data close to i is preferable although this increases the variance and standard error of the estimate. There is, therefore, a trade-off between increased variance at small subsets and bias toward the global coefficients at larger subsets. To address this issue, the bandwidth can be calibrated using the cross validation score (CV) or the corrected Akaike Information Criterion (AICc: see Akaike [1974]). These values express the overall performance of regression models and can take into account the bias-variance trade-off, providing an estimate of the best bandwidth to use [Fotheringham *et al.*, 1998].

[16] A number of statistical tests have been defined to allow the assessment of GWR models against global regression models (for full details, see Fotheringham *et al.* [2002]). These include Analysis of Variance (ANOVA), testing the null hypothesis that the GWR model represents no improvement on the global model, and a Monte Carlo test

to assess the significance of the spatial variability of GWR coefficients. Under the null hypothesis, any random permutation of the data is equally likely to occur.

2.2.3. Application of GWR to Sediment Trap Data

[17] We suggest that geographically weighted regression provides a rigorous approach to assessing the spatial variability of carrying coefficients. Although there has been extensive work on extending the predictive use of GWR [Harris *et al.*, 2011; Kumar and Lal, 2011], GWR is applied here as an exploratory tool. We apply GWR to our updated sediment trap data set using software (GWR 3.0) kindly provided by M. Charlton of the National Centre for Geocomputation at National University of Ireland Maynooth. An adaptive kernel (defining a subset of data by number of nearest neighbors) was chosen to group data because a fixed kernel (defining a subset of data strictly by distance) failed to produce results due to the sparse sampling density in space. A fixed bandwidth could not account for the data points in relative isolation such as those in the Southern Ocean. The AICc minimization technique was used to find the optimal bandwidth. Data were weighted using bi-square function. A Monte Carlo significance test was used to determine whether regression coefficients were spatially variable [Hope, 1968].

3. Results

3.1. Comparison of Global Regression Model Results

3.1.1. Global Regression Analysis

[18] The sediment trap data set shows similar global relationships to those obtained previously with smaller data sets (Figure 1). The correlation between POC and CaCO_3 is strongest ($r = 0.60$), while the correlation between POC and opal is weaker ($r = 0.35$), with lithogenic fluxes being intermediate ($r = 0.45$). Visually, the association of POC with both opal and lithogenic material (Figure 1b) suggests the presence of two separate distributions, one of high POC flux and low mineral flux and the second of low POC flux with high mineral flux (this was originally noted for opal by Klaas and Archer [2002]). The scatterplot for CaCO_3 also displays more variability than previously recognized (Figure 1b). All scatterplots display regional differences when separated into ocean basins as previously noted by Ragueneau *et al.* [2006] and De La Rocha and Passow [2007]. The global POC:mineral ratio is 0.052, close to the ratio observed by Armstrong *et al.* [2001].

[19] Multiple linear regression is used on the global data set to express the flux of POC as a summed linear function of mineral fluxes, and, as observed previously, suggests a dominant role for CaCO_3 (Table 2). The resulting regression model is significant at $p < 0.001$ and predicts 66% of the variability in POC fluxes ($R^2 = 0.66$). The carrying coefficient for CaCO_3 is close to, although very slightly higher than, previous estimates (Table 1). The coefficient for opal is also consistent with previous estimates while the lithogenic coefficient shows the most variability between studies. The lithogenic estimate here is much lower than that found by Klaas and Archer [2002]. Both our estimate and that of Ragueneau *et al.* [2006] are derived from significantly larger data sets ($n = 156\text{--}189$ c.f. $n = 62\text{--}78$) suggesting that this value could be more globally representative.

[20] The R^2 value is relatively low in comparison to the global studies in Table 1. This is due to the inclusion of the

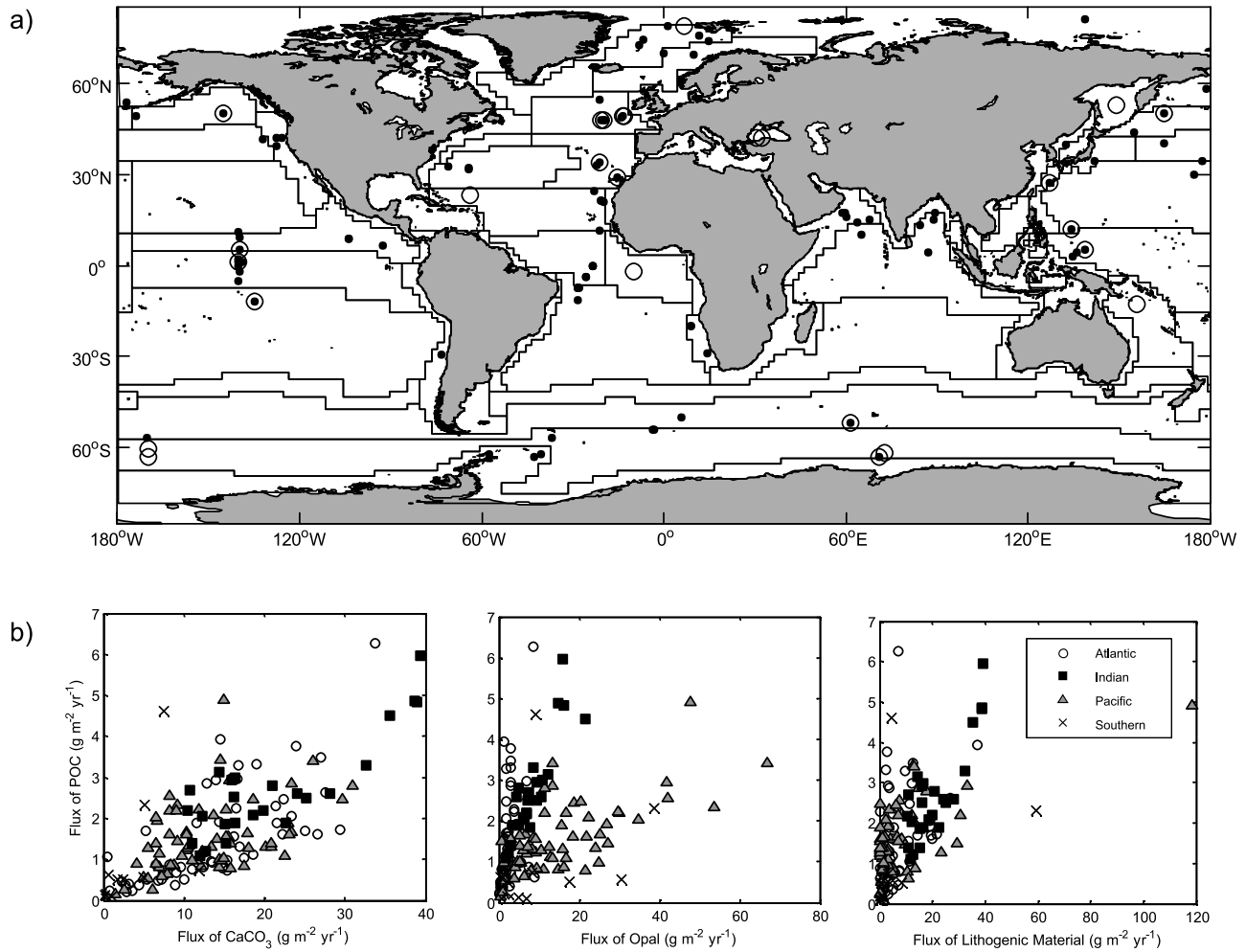


Figure 1. (a) Locations of global sediment traps at >1500 m (dots) and additional data at >1000 m (open circle). Data locations are given in the context of biogeochemical provinces as per *Longhurst* [1998]. (b) Global annual mass fluxes of particulate organic carbon versus global annual mass fluxes of (left) CaCO₃, (middle) opal and (right) lithogenic material as measured by sediment traps >1500 depth. Flux data from different ocean basins are indicated by symbols to highlight regional differences (adapted from *Ragueneau et al.* [2006]).

Table 2. Carrying Coefficients Calculated Using Multiple Linear Regression With Mass Flux Data for Different Depth Ranges and Spatial Scales^a

Global Data	CaCO ₃	Opal	Lithogenic	R ²
>1500 m (n = 156)	0.089	0.023	0.027	0.66
	0.076–0.102	0.012–0.034	0.017–0.037	
>1000 m (n = 186)	0.080	0.017	0.033	0.65
	0.068–0.097	0.007–0.028	0.027–0.039	
Regional Data (>1500 m)				
Atlantic (n = 54)	0.083	0.152 ^a	0.027 ^{NS}	0.58
	0.047–0.118	0.028–0.276	–0.007–0.060	
Indian (n = 25)	0.083	0.058 ^a	0.058	0.94
	0.058–0.108	0.003–0.113	0.034–0.083	
Pacific (n = 63)	0.056	0.033	0.022	0.80
	0.041–0.071	0.025–0.041	0.015–0.028	
Southern Ocean (n = 12)	0.183 ^{NS}	–0.022 ^{NS}	0.034 ^{NS}	0.37 ^{NS}
	–0.065–0.431	–0.139–0.095	–0.050–0.117	

^aAll coefficients are significant at $p < 0.001$ except a where $p < 0.05$. Ninety-five percent confidence intervals are given as $2 \times$ standard error. Note that the model for the Southern Ocean is not significant.

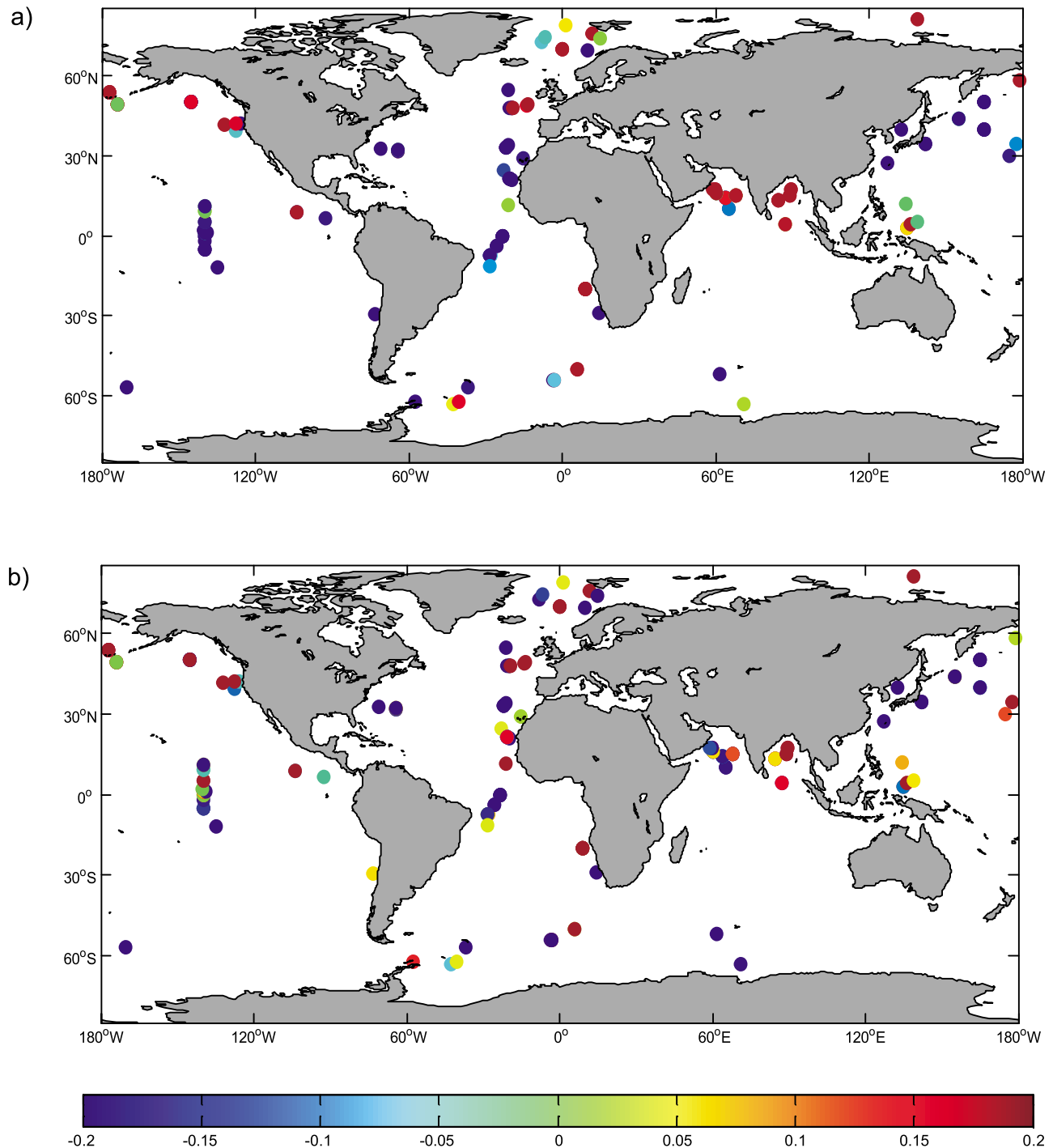


Figure 2. The spatial distribution of residuals from (a) global multiple linear regression and (b) geographically weighted regression analysis. Fluxes of POC are predicted using mass fluxes of CaCO_3 , opal and lithogenic material as in [Klaas and Archer, 2002]. The models predict $\sim 66\%$ and $\sim 82\%$, respectively, of the variability in observed POC fluxes.

intercept term in equation (1). Repeating the analysis without the intercept term increases R^2 to 0.89, a value consistent with previous studies (Table 1). (The large difference between these values is due to the calculation of R^2 from the total sum of squares which is uncorrected without an intercept, leading to inflated values of R^2 [Montgomery *et al.*, 2006].) To be consistent and directly comparable with

geographically weighted regression analysis we include the intercept, justified by the fact the residual mean squares are very similar and that the derived carrying coefficients only differ by a factor of 0.0001. Note, therefore, that R^2 values in this paper are not directly comparable to previous studies.

[21] A regional breakdown of the data to an ocean basin scale is also shown in Table 2 and is comparable to that by

Table 3. Summary Statistics of the GWR Models for Sediment Trap Data at >1500 m^a

Bandwidth	AIC	R ²	ANOVA Sig.	CaCO ₃	Opal	Lithogenic
66 (optimal)	269.82	0.82	<0.001	0.066***	0.033***	0.022***
20	318.04	0.90	<0.001	0.056	0.040***	0.026***
156	313.41	0.71	<0.001	0.089***	0.025	0.026

^aA bandwidth of 66 is the optimal value chosen from the AICc calibration. Additional bandwidths of 20 and 156 are provided to explore the sensitivity of the outcomes to bandwidth. ANOVA statistics test the null hypothesis that the GWR model is no improvement on the global model. 'Global' carrying coefficients are the median of the 156 coefficients. Monte Carlo test for spatial variability rejects the null hypothesis at * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Ragueneau *et al.* [2006] as summarized in Table 1. The carrying coefficient for opal in the Atlantic displays variability compared to other basins and is quite distinct from the global value (0.152 c.f. 0.023). In the Pacific and Indian basins, our spatial coefficients are more consistent with the global values. The Indian basin values differ to those found in Table 1 the reason for which is unclear. The relevant data set in Table 1 is much smaller ($n = 16$ c.f. $n = 25$) and it may be that these coefficients are more influenced by outliers in the regression. In particular, it is also worth considering that the Pacific basin is the largest in size, and potentially includes multiple sources of variability unlike the smaller Indian basin (i.e. as indicated by the number of biogeochemical provinces in each). Therefore, it is uncertain whether the similarity of the Pacific coefficients to global values may be a product of averaging the potentially large spatial variability in fluxes or is actually representative of the values in this area. Table 2 also includes analysis of data from the Southern Ocean but the resulting model is not significant. This highlights the difficulties in conducting regional regression even at the scale of ocean basins. Making comparisons between areas like this is problematic because of varying sample sizes and a lack of consistent statistical methodology.

[22] An alternative method of assessing whether regional variability exists in the global data set is to map the residuals of the global regression model (Figure 2a). An assumption of regression is that the residuals should have a random distribution around zero. Extending this logic, if coefficients are truly global they should also exhibit residuals that are randomly distributed in space [Fotheringham *et al.*, 2002]. However, we note here that negative residuals appear to cluster in the low-latitude Atlantic and Pacific as well as the western sub-arctic Pacific whereas positive residuals cluster in the Arabian Gulf and Indian Ocean. This supports the contention that there is potential non-stationarity in the coefficients which are not truly global.

3.2. Geographically Weighted Regression

[23] Geographically weighted regression analysis was applied to the global sediment trap data set at depths >1500 m. The AICc calibration was used to determine an optimal bandwidth of 66 nearest neighbors.

3.2.1. Assessing the Performance of GWR

[24] A reduction in the AIC score from the global model score of 341.6 and an increase in R^2 from 0.66 to 0.82 suggest that GWR is an improvement on the global model (Table 3). The results of the ANOVA statistics show that the GWR model is a significant improvement on the global model while the results of the Monte Carlo test suggest there

is significant spatial variability in the regression coefficients. A visual comparison of the residuals from the GWR (Figure 2b) against the global regression residuals (Figure 2a) indicates greater heterogeneity in areas previously characterized by clustered residuals, such as the equatorial Pacific, Atlantic and the Indian Ocean. These metrics suggest that the use of a spatially informed regression technique is justified here.

[25] GWR calculates regression coefficients and other statistics at each data point, allowing them to be mapped (Figure 3). The results show distinct regional groupings of the coefficients with minimal variability within these groups. An exception to this are the coefficients for opal in the North Atlantic, which show a range of values in a relatively small area with no identifiable spatial trend. In this region, the bandwidth of 66 nearest neighbors that defines the subsets of data is relatively large in comparison to the area (see Figure 6a) and is therefore influenced by significantly different values in the Arctic and equatorial Atlantic. This highlights a particular issue of using GWR with this data set: the analysis is limited by the relatively small number of data points compared to the area sampled, such that areas of low sampling density may be produce spurious results. Data points at the edge of ocean basins or in sparsely sampled areas, such as the Southern Ocean, will also tend to include data from other basins. The inclusion of data from other ocean basins is an important caveat for this analysis and limits the following section to the discussion of large-scale spatial patterns in coefficients. We hereafter refer to this caveat as inter-basin influence.

[26] To explore the sensitivity of coefficients to the bandwidth and inter-basin influence, the bandwidth was manually changed to 20 and 156 nearest neighbors in comparison to the calibrated 66 neighbors (Table 3), and the results from the most variable mineral, CaCO₃, were plotted (Figures 4a and 4b). A bandwidth of 156 is used to assess if the GWR technique can recover MLRA global coefficient values. Although this will include all data points in the data set they will still be weighted by distance. Figure 4b shows that most of the coefficients converge toward a global value around 0.09. The equatorial Pacific is the only area not to conform to the global value. This is likely to be a result of a combination of densely sampled data being weighted heavily but in relative isolation to other data points. The higher AIC score suggests this model has less ability to predict POC than the calibrated optimal model, which is also suggested by the lower R^2 value (0.71) although it is still higher than the global mean MLRA model (0.66) (Table 1). With a much smaller bandwidth of 20, the R^2 value increases but the higher AIC score indicates other aspects of the

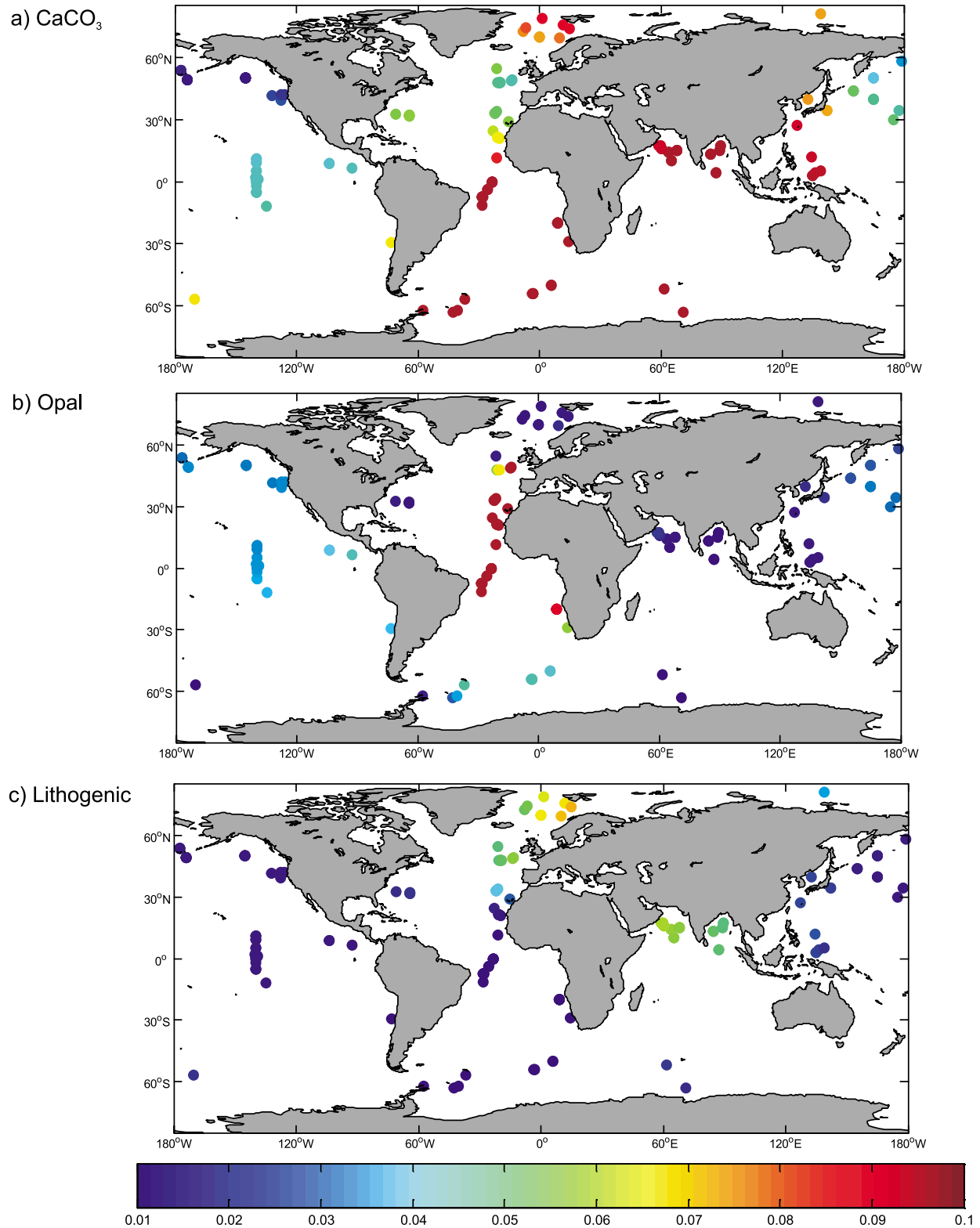


Figure 3. The spatial distribution of coefficients calculated using geographically weighted regression analysis for (a) CaCO_3 , (b) opal, and (c) lithogenic material. The model is the same multiple linear regression model applied to the global data in Table 2 and Figure 2. The GWR analysis uses a bandwidth of 66 nearest neighbors, defined by an AICc minimization calibration procedure. A bi-square weighting scheme was used. The GWR model predicts $\sim 82\%$ of the variability in POC fluxes and is an improvement on the global model. The corresponding global coefficient values are 0.089, 0.023 and 0.027 for (a), (b) and (c), respectively.

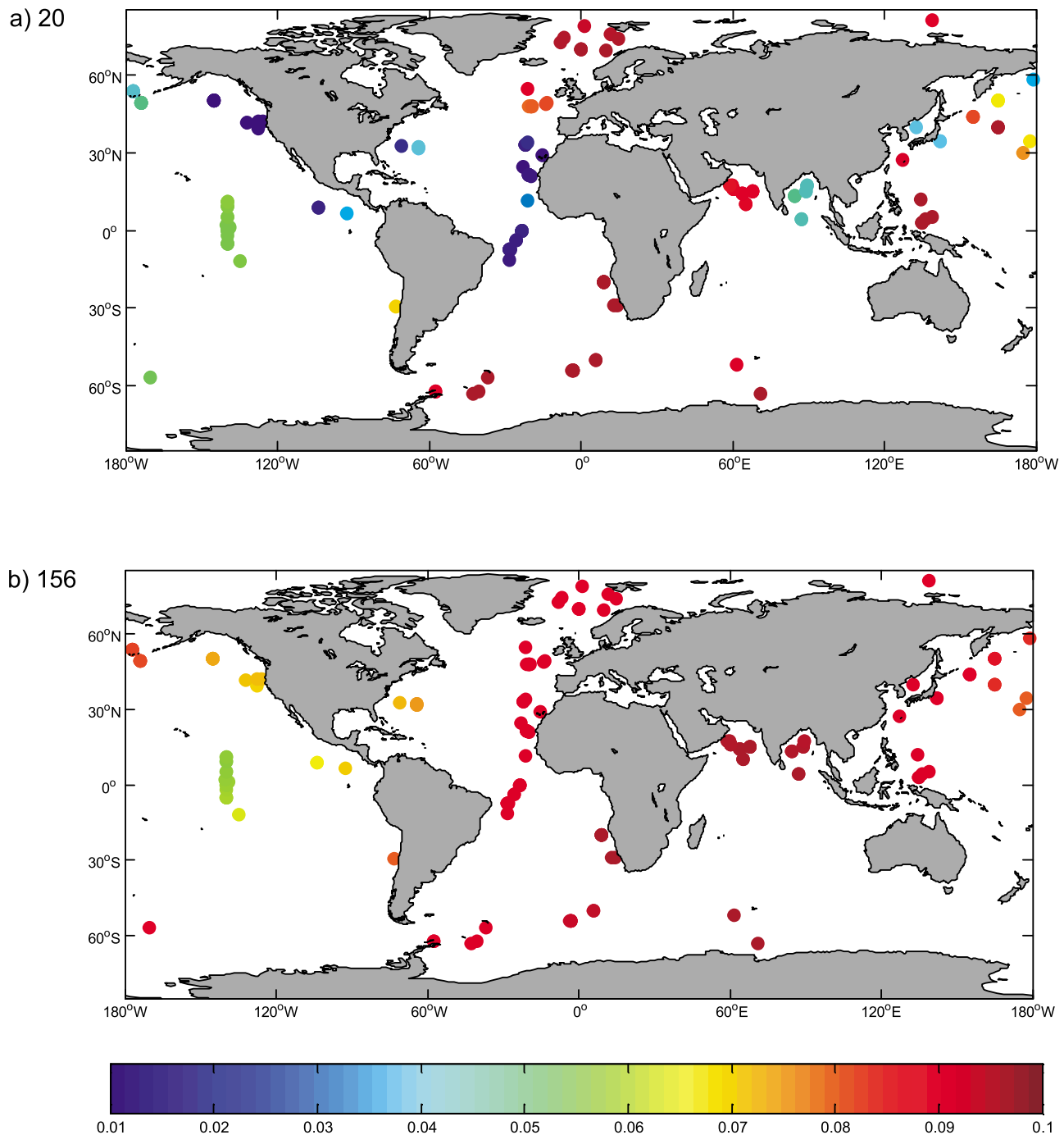


Figure 4. Spatial distribution for CaCO_3 :POC coefficients calculated using geographically weighted regression as in Figure 3 but with manually altered bandwidth values of (a) 20 and (b) 156 nearest neighbors. Only coefficients for CaCO_3 are displayed, as it displays the most variability of the three minerals considered. For reference, the global CaCO_3 coefficient is 0.089.

model are worse, such as increased variance (Table 3). We might expect coefficients to vary significantly at this bandwidth because the coefficients represent local values but are also subject to influence from outliers. However, the large-scale general spatial patterns (Figure 4a) are comparable to those in Figure 3a, suggesting these are not an artifact of bandwidth size. A notable exception to this is the lower coefficients in the equatorial Atlantic (Figure 4a, cf.

Figure 3a). This suggests that the subset of data defined by a bandwidth of 66 nearest neighbors may be influenced by significantly higher values in the Southern Ocean but overall our ability to capture the global mean MLRA coefficient values when relaxing the spatial bandwidth gives us increased confidence in this method.

[27] The basin-scale coefficients in Table 2 appear to corroborate the GWR coefficients. Furthermore, to test the

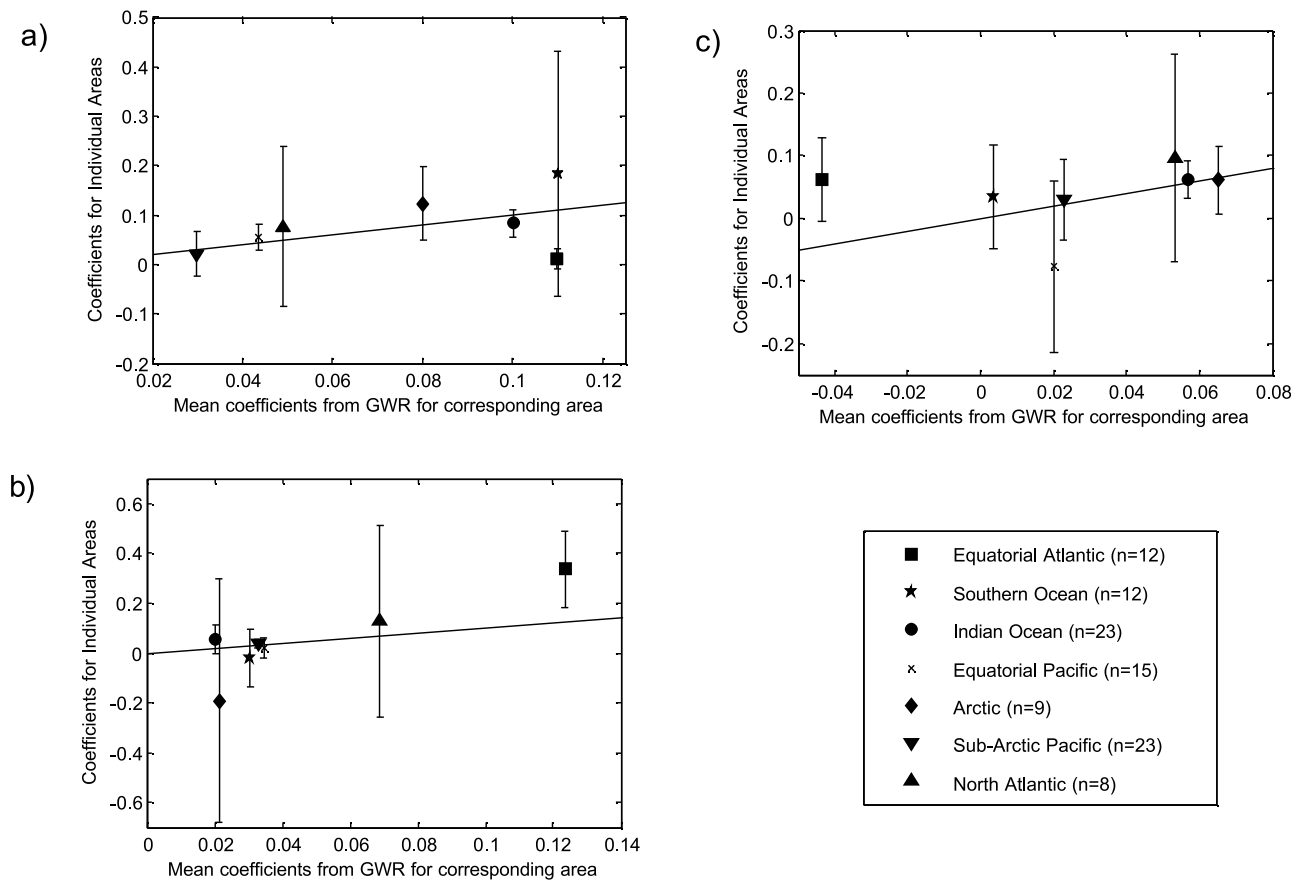


Figure 5. Comparison of regression coefficients calculated over manually selected areas and mean coefficient values calculated using geographically weighted regression for the corresponding area for (a) CaCO_3 , (b) opal, and (c) lithogenic material. Errorbars correspond to $2 \times$ standard error; 1:1 ratio lines are shown.

validity of large scale patterns of coefficients produced by GWR we manually selected subsets of data by region and compared them to the mean GWR coefficients for the corresponding area (Figure 5). Overall there is general agreement between the coefficients from both methods, supporting the results of the GWR analysis. This method of validation is difficult because some areas have a small number of data points, which can lead to statistically unreliable outcomes and is one reason why GWR is a preferable technique. Overall, the coefficients show general agreement with the GWR coefficients, suggesting that the GWR technique is producing reliable results even when considering the potential for inter-basin influence.

3.2.2. Spatial Patterns in Regression Coefficients

[28] The geographically weighted regression analysis defines clear regional patterns in the coefficients for CaCO_3 , opal and lithogenic particles (Figure 3). The Atlantic displays some of the more unexpected results, with coefficients differing appreciably from global values. Coefficients for both CaCO_3 and opal display a decreasing trend with increasing latitude between 30 and 60 degrees North in the North Atlantic. Opal is quantitatively more important than CaCO_3 and lithogenics in the low latitude Atlantic, corroborating the basin-scale approach in Tables 1 and 2. In the

Arctic Ocean however, CaCO_3 and lithogenics are quantitatively more important. This is not revealed by the basin-scale results.

[29] Overall in the Pacific, there is much less difference between coefficient values, with no mineral showing overall importance except for the higher coefficients observed for CaCO_3 in the western and equatorial Pacific. These could be a result of inter-basin influence from the Indian Ocean although similar behavior is not observed in the lithogenic coefficients, suggesting that it is probably not an artifact. The majority of the CaCO_3 coefficients in the Pacific are much lower than the global value of 0.089 (as also highlighted in the basin-scale analysis) (Table 2).

[30] The Indian Ocean is one of the only regions that display similar coefficients to those from the global analysis, with a quantitative importance identified between both CaCO_3 and lithogenics, and POC. Variability between the Arabian Gulf and Bay of Bengal is small and is likely a result of subtle changes in the subsets of data used. Finally, the Southern Ocean also displays unexpected results, with quantitative importance between CaCO_3 and POC only. Inter-basin influence may be a large problem for the Southern Ocean, given the small number of samples and the relative distances between them (see Figure 6a for indication

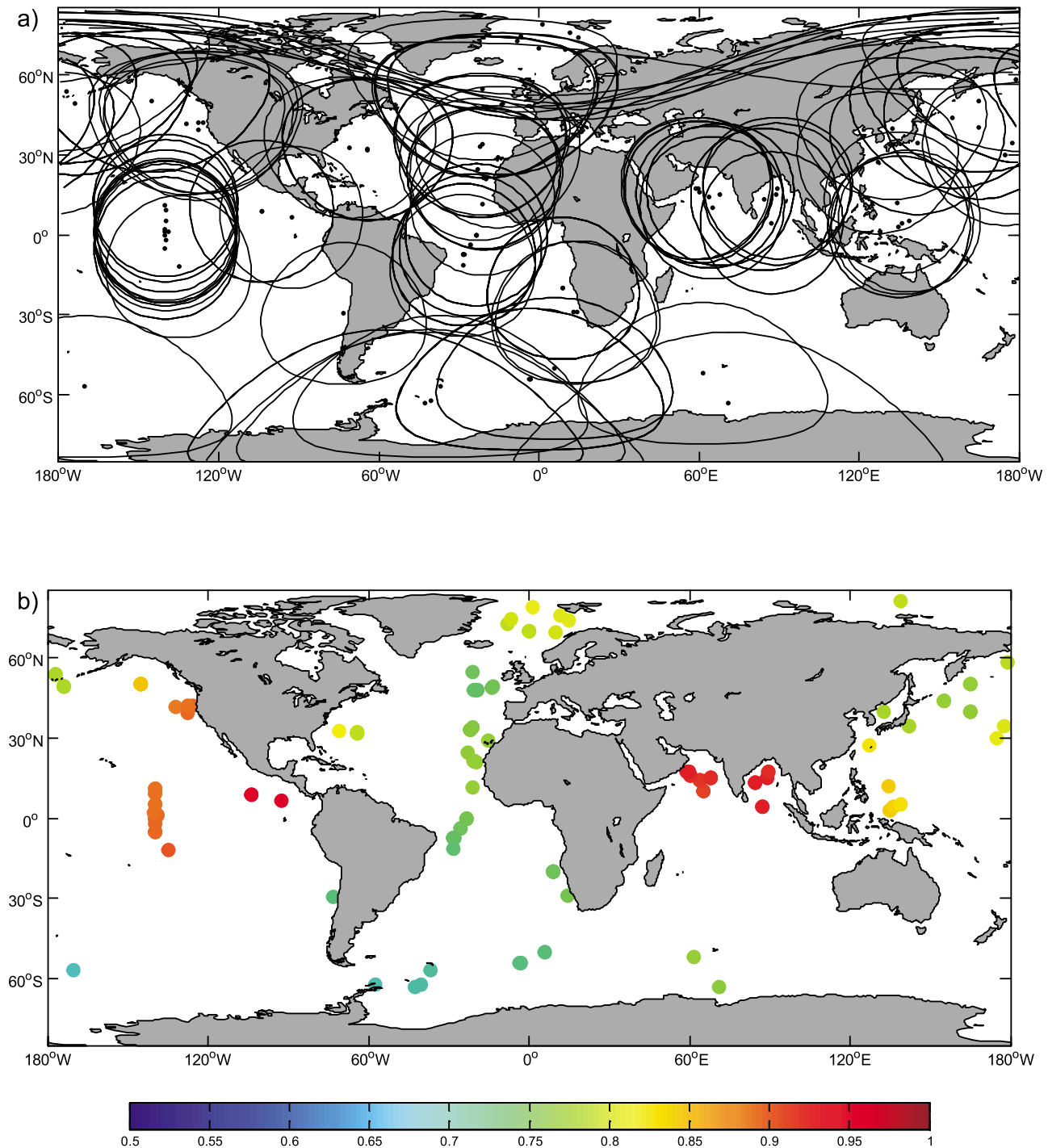


Figure 6. a) Schematic depicting the bandwidth of geographically weighted regression. Circular line indicates distance at which a weight of 0.75 occurs. Data is weighted ≥ 0.75 within the area and < 0.75 up to the distance of the last nearest neighbor. A bandwidth of 66 nearest neighbors is used. (b) Local R^2 values from the geographically weighted regression in Figure 3 using a bandwidth of 66 nearest neighbors and bi-square weighting scheme.

of bandwidth sizes). This is reflected in both the insignificant regression result in Table 2 and the relatively low R^2 values generated by the GWR analysis (Figure 6b).

[31] Overall the global coefficient for CaCO_3 of 0.089 appears to be dominated by the Southern Ocean, the Indian

Ocean, western equatorial Pacific and the Arctic while regionally displaying much lower values (< 0.04 – 0.05) in the Atlantic and the remaining Pacific. In contrast, the global value of 0.023 for opal and 0.027 for lithogenics appear more consistent over large regions, with exceptions in the

Atlantic for opal and the Arctic and Indian Oceans for lithogenics.

4. Discussion

[32] Geographically weighted regression (GWR) offers a promising technique to explore the regional variability of regression coefficients. However, its application to sediment trap data is limited both by the amount of data available and the geographical distribution of sampling, which is somewhat clustered, being a collection of individual projects focused on specific areas and research questions. The overall size of the data set is also at the lower end to which GWR may be successfully applied (M. E. Charlton, personal communication, 2012). GWR is also subject to some of the caveats to previous global mean (and annual average) analyses and which may contribute to the observed variability. For instance, differences in the origin of sinking particles could have an influence on the degree of coupling between POC and ballast mineral. Foraminiferal CaCO_3 can be de-coupled from POC due to the reproductive cycle, where the gametes (contributing POC) abandon the organism, leaving particles that are predominantly CaCO_3 to sink [Loubere *et al.*, 2007]. Ziveri *et al.* [2007] showed that the relationship between POC and CaCO_3 in different species of coccolithophores can also be variable, leading to differences in regression coefficients. Similarly, Thunell *et al.* [2007] observed variable POC to opal ratios in different coastal diatom species which contributed to variation in regression coefficients. Because sinking particles captured by sediment traps are expressed as bulk CaCO_3 and opal, it is likely this could be a source of variability in the data which cannot be accounted for here. The method of estimating lithogenic fluxes has also been highlighted as an issue in analyzing fluxes [Boyd and Trull, 2007]. Defined as the difference between total mass flux and the main biogenic components, the lithogenic fraction can be more accurately termed the residual flux [Salter *et al.*, 2010] and may include material that is unaccounted for. Unfortunately this is not something we have been able to address in this study.

[33] Despite these caveats, GWR removes some of the subjectivity in choosing subsets of data and provides a framework for dealing with statistical problems, e.g. such as increased variance of coefficients. GWR therefore offers a more objective technique and one able to explore spatial patterns at a finer resolution than the basin-scale analyses previously possible.

4.1. Why is there Spatial Variability in Carrying Coefficients?

[34] If there was a systematic mechanism linking increased sinking velocity of particulate organic matter with CaCO_3 or between the efficiency of export to depth and protection from remineralization [Armstrong *et al.*, 2001; Klaas and Archer, 2002], we would expect to observe consistent relationships between POC and the presence of minerals at a global scale. However, in the GWR analysis used here we have identified significant spatial variability in the carrying coefficients but with generally coherent large-scale spatial relationships. If we are to improve models of the ocean carbon cycle and hence better quantify feedbacks and potential impacts of global change, we need to understand why any coherent

patterns occur at all and why regions differ from one another. The underlying complexity and challenge in mechanistic interpretation is illustrated by considering the quantitative importance observed for opal in the Atlantic and CaCO_3 in the Southern Ocean, which is somewhat counterintuitive considering the global distribution of productivity by major plankton groups, with relatively little opal exported from the surface of the Atlantic and relatively little CaCO_3 exported in the Southern Ocean. The inference is that fluxes of POC are quantitatively associated with fluxes of opal in the Atlantic despite this being an area of very low opal fluxes. The inverse applies to the Southern Ocean and CaCO_3 . This suggests that there is more to the controls of sinking POC than just the dominant mineral phase being produced at the surface.

[35] Multiple linear regression reflects the joint variability between all three minerals and POC and therefore coefficients reflect variability related to the specific combination of all flux components and should not be interpreted in terms of any one mineral component in isolation. Honjo *et al.* [2008] demonstrated that the relationships between the proportions of POC, CaCO_3 and opal within mass fluxes are broadly aligned with regional patterns such as biogeochemical provinces. They used this to define broad regions that are dominated either by POC, CaCO_3 or opal. The patterns of coefficients in Figure 3 define broad regions and appear consistent with this interpretation, being visually comparable to biogeochemical provinces (and which we have demonstrated are not artifacts of the statistical method or inter-basin influence; see Figure 5). The regional regression coefficients we observe here may then reveal variability derived from the specific combinations of flux components that ultimately derive from specific ecosystem characteristics.

[36] Recent studies have built on this alternative interpretation of the observed relationships between POC and minerals, focusing on the variability originating from ecosystem processes. Francois *et al.* [2002] originally suggested that differences in the biodegradability of POC derived from CaCO_3 -dominated ecosystems resulted in efficient POC transfer rather than from the direct presence of CaCO_3 itself. Developing this methodology further with satellite data, Henson *et al.* [2012] found that the export efficiency of particles from the surface is low in carbonate dominated regions but transfer efficiency of POC to depth is high, with the reverse being true in regions dominated by diatoms. This complements similar findings in the Southern Ocean [Lam and Bishop, 2007] and the identification of similar variability through the interaction between the timing and intensity of activity in producer and consumer communities [Lam *et al.*, 2011]. Our results are consistent with the relationships between mass flux components being derived from ecosystem characteristics, and support the findings of these studies in suggesting that the key factor in the variability of POC reaching the deep ocean is the interaction between producer and consumer communities in different ecosystems, possibly reflecting the resulting differences in exported organic matter biodegradability. In ecosystems with highly seasonal productivity, mismatches in the timing of producer and consumer activity can result in labile organic matter being exported which is then readily remineralized in the water column. Likewise, in non-seasonal ecosystems more constant production and consumption can result in the export of relatively refractory material which is

more resistant to remineralization in the water column and sinking to depth [Lam *et al.*, 2011]. Our spatial analysis suggests that even within diatom and carbonate dominated regions there are distinct variations which may relate to specific differences in this packaging function, such as the distribution of different species and behavior of zooplankton present, and the subsequent action of bacteria [Lam and Bishop, 2007; Buesseler and Boyd, 2009].

4.2. Implications for the Ballast Hypothesis and Modeling

[37] The ballast hypothesis has been highly influential in proposing that fluxes of POC can be mechanistically linked to fluxes of ballast minerals such as CaCO_3 . The basis of this hypothesis is the observed strong global quantitative relationship observed between POC and CaCO_3 . However, we have shown that this strong relationship does not apply to all regions (or even sub-regions) in the ocean and is not constrained to CaCO_3 alone. In effect, the strong relationship appears to be an artifact of averaging fluxes on a global scale, masking important regional variability. This could occur because globally there is a large range of flux magnitudes in both POC and CaCO_3 which is larger than the variability of the regional POC: CaCO_3 ratios. This is evident in Figure 1b, where different basin scale ratios can be seen within the global data distribution. Using the relative standard deviation (RSD) as a comparable measure of the variation within each data set, the individual fluxes of CaCO_3 and POC have RSD values of 0.62 and 0.71 respectively whereas the RSD of the CaCO_3 carrying coefficients is smaller at 0.45. As such, we suggest the appearance of a global correlation between CaCO_3 and POC could be exaggerated as a result of considering data at a global scale. Equally, this may bias the interpretation of such relationships without information on spatial variability. In a similar manner, Lam *et al.* [2011] produced global coefficient values when averaging particle concentration data through time but which masked important temporal variability, suggesting a different mechanistic interpretation. The analyses of temporal variability [Lam *et al.*, 2011], transfer efficiencies of POC [Francois *et al.*, 2002; Henson *et al.*, 2012] and the regional variability highlighted in this study all suggest that ecosystem-based mechanisms are influential in setting the efficiency of the biological pump. While this does not rule out ballasting as a mechanism (e.g. Ploug *et al.* [2008]; Engel *et al.* [2009]; Iversen and Ploug [2010]) it shifts the focus away from a simple causal physical explanation.

[38] The ballast hypothesis has inspired specific parameterizations using the global statistical coefficients to be adopted in a number of global ocean carbon cycle models (e.g. HAMOCC5.1 [Howard *et al.*, 2006], CCSR COCO 4.0 [Oka *et al.*, 2008] and POTSDAM-C [Hofmann and Schellnhuber, 2009], see also PISCES [Gehlen *et al.*, 2006]). Our results suggest that using these global statistical parameters explicitly in models requires careful reconsideration. However, moving away from a common global mechanism to parameterizations able to capture spatial variability in the flux relationships is not trivial. Application of a simple prescribed map (distribution) of carrying capacities, while potentially improving the simulation of dissolved nutrient and carbon distributions in the ocean interior, is likely to fail to provide an appropriate response to global change.

Instead, there needs to be a shift in the focus away from a geochemical-based understanding and parameterization approaches toward a more ecological-based understanding of fluxes [Ragueneau *et al.*, 2006; Boyd and Trull, 2007], which will require alternative mechanistic representations in models. Reflecting ecosystem function will then lead to potentially very different feedback mechanisms when considering ocean acidification and climate change [Henson *et al.*, 2012]. Further work will be needed to develop flux parameterizations that link aspects of ecosystem function with organic matter particle fluxes. In this respect, Riley *et al.* [2012] recently suggested a model that considers suspended particles, slow sinking particles that are subject to remineralization in the water column, and fast sinking particles that may be subject to ballasting. Considering the range of mechanistic interpretations and inherent uncertainty this introduces, further work is needed to quantitatively constrain the range of feedback processes through modeling. This may offer a new method of evaluation when applied to events in the paleo-record, which may be analogous to current ocean acidification.

5. Conclusions

[39] Geographically weighted regression (GWR) has allowed us to explore the spatial variability in the particulate organic carbon carrying coefficients associated with ballast mineral fluxes in the deep ocean. GWR reveals significant spatial variability in the relationship between POC and minerals that is obscured when using a global multiple linear regression analysis (MLRA) approach. This leads us to conclude that the strong and apparently near-uniform relationship between POC and CaCO_3 observed at a global scale is an artifact of high degree of spatial averaging. Regional MLRA goes some way to resolving the spatial variability that exists in coefficients, but still hides significant spatial variability within large regions.

[40] The nature of the spatial patterns in carrying coefficients observed using GWR may result from unique combinations of flux components reaching depth which are in turn derived from specific properties of the surface ecosystem. This has important implications for modeling studies which parameterize a ballasting mechanism based directly on global MLRA. New parameterizations (and understanding) need to be developed that link aspects of ecosystem function with organic matter particle fluxes. As there are multiple hypotheses that may explain the observations and patterns of correlations between fluxes, new models will be needed that can explore different mechanisms and quantitatively evaluate their implications against observations and ultimately to quantify the nature and strength of feedbacks in the system.

[41] **Acknowledgments.** This work was conducted as part of a project studentship (JW) associated with the UK Ocean Acidification Research Programme (UKOARP) grant NE/H017240/1 to AR and SB. AR acknowledges support from a Royal Society University Research Fellowship. We thank two anonymous reviewers for their helpful comments.

References

Akaike, H. (1974), A new look at the statistical model identification, *IEEE Trans. Autom. Control*, 19(6), 716–723, doi:10.1109/TAC.1974.1100705.

- Archer, D., and E. Maier-Reimer (1994), Effect of deep-sea sedimentary calcite preservation on atmospheric CO₂ concentration, *Nature*, 367(6460), 260–263, doi:10.1038/367260a0.
- Armstrong, R. A., C. Lee, J. I. Hedges, S. Honjo, and S. G. Wakeham (2001), A new, mechanistic model for organic carbon fluxes in the ocean based on the quantitative association of POC with ballast minerals, *Deep Sea Res., Part II*, 49(1–3), 219–236.
- Barker, S., J. A. Higgins, and H. Elderfield (2003), The future of the carbon cycle: Review, calcification response, ballast and feedback on atmospheric CO₂, *Philos. Trans. R. Soc. A*, 361(1810), 1977–1999, doi:10.1098/rsta.2003.1238.
- Boyd, P. W., and T. W. Trull (2007), Understanding the export of biogenic particles in oceanic waters: Is there consensus?, *Prog. Oceanogr.*, 72(4), 276–312, doi:10.1016/j.pocean.2006.10.007.
- Brunsdon, C., A. S. Fotheringham, and M. E. Charlton (1998), Geographically weighted regression-modelling spatial non-stationarity, *Statistician*, 47(3), 276–312, doi:10.1111/1467-9884.00145.
- Buesseler, K., and P. Boyd (2009), Shedding light on processes that control particle export and flux attenuation in the twilight zone of the open ocean, *Limnol. Oceanogr.*, 54(4), 1210–1232, doi:10.4319/lo.2009.54.4.1210.
- Conte, M., R. Ralph, and E. Ross (2001), Seasonal and interannual variability in deep ocean particle fluxes at the Ocean Flux Program (OFP)/Bermuda Atlantic Time Series (BATS) site in the western Sargasso Sea near Bermuda, *Deep Sea Res., Part II*, 48(8–9), 1471–1505, doi:10.1016/S0967-0645(00)00150-8.
- De La Rocha, C. L., and U. Passow (2007), Factors influencing the sinking of POC and the efficiency of the biological carbon pump, *Deep Sea Res., Part II*, 54(5–7), 639–658, doi:10.1016/j.dsr2.2007.01.004.
- De La Rocha, C. L., N. Nowald, and U. Passow (2008), Interactions between diatom aggregates, minerals, particulate organic carbon, and dissolved organic matter: Further implications for the ballast hypothesis, *Global Biogeochem. Cycles*, 22, GB4005, doi:10.1029/2007GB003156.
- Engel, A., J. Szlosek, L. Abramson, Z. F. Liu, and C. Lee (2009), Investigating the effect of ballasting by CaCO₃ in *Emiliania huxleyi*: I. Formation, settling velocities and physical properties of aggregates, *Deep Sea Res., Part II*, 56(18), 1396–1407, doi:10.1016/j.dsr2.2008.11.027.
- Fotheringham, A. S., M. E. Charlton, and C. Brunsdon (1998), Geographically weighted regression: A natural evolution of the expansion method for spatial data analysis, *Environ. Plan. A*, 30(11), 1905–1927, doi:10.1068/a301905.
- Fotheringham, A. S., C. Brunsdon, and M. E. Charlton (2002), *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*, Wiley, Chichester, U. K.
- Francois, R., S. Honjo, R. Krishfield, and S. Manganini (2002), Factors controlling the flux of organic carbon to the bathypelagic zone of the ocean, *Global Biogeochem. Cycles*, 16(4), 1087, doi:10.1029/2001GB001722.
- Gehlen, M., L. Bopp, N. Ernprin, O. Aumont, C. Heinze, and O. Ragueneau (2006), Reconciling surface ocean productivity, export fluxes and sediment composition in a global biogeochemical ocean model, *Biogeosciences*, 3(4), 521–537.
- Harris, P., C. Brunsdon, and A. S. Fotheringham (2011), Links, comparisons and extensions of the geographically weighted regression model when used as a spatial predictor, *Stoch. Environ. Res. Risk Assess.*, 25(2), 123–138, doi:10.1007/s00477-010-0444-6.
- Heinze, C. (2004), Simulating oceanic CaCO₃ export production in the greenhouse, *Geophys. Res. Lett.*, 31, L16308, doi:10.1029/2004GL020613.
- Henson, S., R. Sanders, and E. Madsen (2012), Global patterns in efficiency of particulate organic carbon export and transfer to the deep ocean, *Global Biogeochem. Cycles*, 26, GB1028, doi:10.1029/2011GB004099.
- Hofmann, M., and H. J. Schellnhuber (2009), Oceanic acidification affects marine carbon pump and triggers extended marine oxygen holes, *Proc. Natl. Acad. Sci. U. S. A.*, 106(9), 3017–3022, doi:10.1073/pnas.0813384106.
- Honda, M. C., and S. Watanabe (2010), Importance of biogenic opal as ballast of particulate organic carbon (POC) transport and existence of mineral ballast-associated and residual POC in the Western Pacific Subarctic Gyre, *Geophys. Res. Lett.*, 37, L02605, doi:10.1029/2009GL041521.
- Honjo, S., S. J. Manganini, R. A. Krishfield, and R. Francois (2008), Particulate organic carbon fluxes to the ocean interior and factors controlling the biological pump: A synthesis of global sediment trap programs since 1983, *Prog. Oceanogr.*, 76(3), 217–285, doi:10.1016/j.pocean.2007.11.003.
- Hope, A. C. A. (1968), A simplified Monte Carlo significance test procedure, *J. R. Stat. Soc. B*, 30(3), 582–598.
- Howard, M. T., A. M. E. Winguth, C. Klaas, and E. Maier-Reimer (2006), Sensitivity of ocean carbon tracer distributions to particulate organic flux parameterizations, *Global Biogeochem. Cycles*, 20, GB3011, doi:10.1029/2005GB002499.
- Iversen, M. H., and H. Ploug (2010), Ballast minerals and the sinking carbon flux in the ocean: Carbon-specific respiration rates and sinking velocity of marine snow aggregates, *Biogeosciences*, 7(9), 2613–2624, doi:10.5194/bg-7-2613-2010.
- Klaas, C., and D. E. Archer (2002), Association of sinking organic matter with various types of mineral ballast in the deep sea: Implications for the rain ratio, *Global Biogeochem. Cycles*, 16(4), 1116, doi:10.1029/2001GB001765.
- Kumar, S., and R. Lal (2011), Mapping the organic carbon stocks of surface soils using spatial interpolator, *J. Environ. Monitor.*, 13, 3128–3135, doi:10.1039/c1em10520e.
- Lam, P. J., and J. K. B. Bishop (2007), High biomass, low export regimes in the Southern Ocean, *Deep Sea Res., Part II*, 54(5–7), 601–638, doi:10.1016/j.dsr2.2007.01.013.
- Lam, P. J., S. C. Doney, and J. K. B. Bishop (2011), The dynamic ocean biological pump: Insights from a global compilation of particulate organic carbon, CaCO₃, and opal concentration profiles from the mesopelagic, *Global Biogeochem. Cycles*, 25, GB3009, doi:10.1029/2010GB003868.
- Longhurst, A. (1998), *Ecological Geography of the Sea*, Acad. Press, San Diego, Calif.
- Loubere, P., S. A. Siedlecki, and L. I. Bradtmiller (2007), Organic carbon and carbonate fluxes: Links to climate change, *Deep Sea Res., Part II*, 54(5–7), 437–446, doi:10.1016/j.dsr2.2007.02.001.
- Montgomery, D. C., E. A. Peck, and G. G. Vining (2006), *Introduction to Linear Regression Analysis*, 4th ed., Wiley, Hoboken, N. J.
- Mortlock, R., and P. Froelich (1989), A simple method for the rapid determination of biogenic opal in pelagic marine sediments, *Deep Sea Res., Part A*, 36(9), 1415–1426, doi:10.1016/0198-0149(89)90092-7.
- Oka, A., S. Kato, and H. Hasumi (2008), Evaluating effect of ballast mineral on deep-ocean nutrient concentration by using an ocean general circulation model, *Global Biogeochem. Cycles*, 22, GB3004, doi:10.1029/2007GB003067.
- Passow, U. (2004), Switching perspectives: Do mineral fluxes determine particulate organic carbon fluxes or vice versa?, *Geochem. Geophys. Geosyst.*, 5, Q04002, doi:10.1029/2003GC000670.
- Passow, U., and C. L. De La Rocha (2006), Accumulation of mineral ballast on organic aggregates, *Global Biogeochem. Cycles*, 20, GB1013, doi:10.1029/2005GB002579.
- Ploug, H., M. H. Iversen, and G. Fischer (2008), Ballast, sinking velocity, and apparent diffusivity within marine snow and zooplankton fecal pellets: Implications for substrate turnover by attached bacteria, *Limnol. and Oceanogr.*, 53(5), 1878–1886, doi:10.4319/lo.2008.53.5.1878.
- Ragueneau, O., S. Schultes, K. Bidle, P. Claquin, and B. La Moriceau (2006), Si and C interactions in the world ocean: Importance of ecological processes and implications for the role of diatoms in the biological pump, *Global Biogeochem. Cycles*, 20, GB4S02, doi:10.1029/2006GB002688.
- Ridgwell, A. J. (2003), An end to the “rain ratio” reign?, *Geochem. Geophys. Geosyst.*, 4, 1051, doi:10.1029/2003GC000512.
- Riebesell, U., A. Kortzinger, and A. Oschlies (2009), Sensitivities of marine carbon fluxes to ocean change, *Proc. Natl. Acad. Sci. U. S. A.*, 106(49), 20,602–20,609, doi:10.1073/pnas.0813291106.
- Riley, J., R. Sanders, F. Marsay, C. Moigne, E. Achterberg, and A. Poulton (2012), The relative contribution of fast and slow sinking particles to ocean carbon export, *Global Biogeochem. Cycles*, 26, GB1026, doi:10.1029/2011GB004085.
- Salter, I., A. Kemp, R. Lampitt, and M. Gledhill (2010), The association between biogenic and inorganic minerals and the amino acid composition of settling particles, *Limnol. Oceanogr.*, 55, 2207–2218, doi:10.4319/lo.2010.55.5.2207.
- Scholten, J. C., et al. (2001), Trapping efficiencies of sediment traps from the deep eastern North Atlantic: The 230th calibration, *Deep Sea Res., Part II*, 48(10), 2383–2408, doi:10.1016/S0967-0645(00)00176-4.
- Stemann, L., G. Jackson, and G. Gorsky (2004), A vertical model of particle size distributions and fluxes in the midwater column that includes biological and physical process—Part II: Application to a three year survey in the NW Mediterranean Sea, *Deep Sea Res., Part I*, 51(7), 885–908, doi:10.1016/j.dsr.2004.03.002.
- Thunell, R., C. Benitez-Nelson, R. Varela, Y. Astor, and F. Muller-Karger (2007), Particulate organic carbon fluxes along upwelling-dominated continental margins: Rates and mechanisms, *Global Biogeochem. Cycles*, 21, GB1022, doi:10.1029/2006GB002793.
- Vichi, M., J. I. Allen, S. Masina, and N. J. Hardman-Mountford (2011), The emergence of ocean biogeochemical provinces: A quantitative assessment and a diagnostic for model evaluation, *Global Biogeochem. Cycles*, 25, GB2005, doi:10.1029/2010GB003867.
- Wong, C., F. Whitney, K. Crawford, K. Iseki, R. Matear, W. Johnson, J. Page, and D. Timothy (1999), Seasonal and interannual variability in particle fluxes of carbon, nitrogen and silicon from time series of sediment traps at Ocean Station P, 1982–1993: Relationship to changes in subarctic

- primary productivity, *Deep Sea Res., Part II*, 46(11–12), 2735–2760, doi:10.1016/S0967-0645(99)00082-X.
- Yu, E. F., R. Francois, M. P. Bacon, S. Honjo, A. P. Fleer, S. J. Manganini, M. M. R. van der Loeff, and V. Ittekkot (2001), Trapping efficiency of bottom-tethered sediment traps estimated from the intercepted fluxes of ^{230}Th and ^{231}Pa , *Deep Sea Res., Part I*, 48(3), 865–889, doi:10.1016/S0967-0637(00)00067-4.
- Ziveri, P., B. de Bernardi, K. H. Baumann, H. M. Stoll, and P. G. Mortyn (2007), Sinking of coccolith carbonate and potential contribution to organic carbon ballasting in the deep ocean, *Deep Sea Res., Part II*, 54(5–7), 659–675, doi:10.1016/j.dsr2.2007.01.006.